University Library



A gateway to Melbourne's research publications

Minerva Access is the Institutional Repository of The University of Melbourne

Author/s:
Raghunathan, S;Holder, GP;Bartlett, JG;Patil, S;Reichardt, CL;Whitehorn, N

Title:
An inpainting approach to tackle the kinematic and thermal SZ induced biases in CMB-cluster lensing estimators

Date:
2019-11-26

Citation:
Raghunathan, S., Holder, G. P., Bartlett, J. G., Patil, S., Reichardt, C. L. & Whitehorn, N. (2019). An inpainting approach to tackle the kinematic and thermal SZ induced biases in CMB-cluster lensing estimators. Journal of Cosmology and Astroparticle Physics, 2019 (11), https://doi.org/10.1088/1475-7516/2019/11/037.

Persistent Link:
https://hdl.handle.net/11343/251357

Accepted for publication in JCAP

arXiv:1904.13392v2 [astro-ph.CO] 7 Nov 2019

An Inpainting Approach to Tackle the Kinematic and Thermal SZ Induced Biases in CMB-Cluster Lensing Estimators

Srinivasan Raghunathan[a] **Gilbert P. Holder**[b] **James G. Bartlett**[c,d] **SanjayKumar Patil**[e] **Christian L. Reichardt**[e] **Nathan Whitehorn**[a]

[a]Department of Physics and Astronomy, University of California, Los Angeles, CA, USA 90095

[b]Department of Astronomy and Department of Physics, University of Illinois, 1002 West Green St., Urbana, IL, USA 61801

[c]APC, AstroParticule et Cosmologie, Université Paris Diderot, CNRS/IN2P3, CEA/lrfu, Observatoire de Paris, Sorbonne Paris Cité, 10, rue Alice Domon et Léonie Duquet, Paris Cedex 13, France

[d]Jet Propulsion Laboratory, California Institute of Technology, 4800 Oak Grove Drive, Pasadena, CA, USA 91109

[e]School of Physics, University of Melbourne, Parkville, VIC 3010, Australia

E-mail: sri@physics.ucla.edu, gholder@illinois.edu, bartlett@apc.univ-paris7.fr, s.patil2@student.unimelb.edu.au, christian.reichardt@unimelb.edu.au, nwhitehorn@physics.ucla.edu

Abstract. A galaxy cluster's own Sunyaev-Zel'dovich (SZ) signal is known to be a major contaminant when reconstructing the cluster's underlying lensing potential using cosmic microwave background (CMB) temperature maps. In this work, we develop a modified quadratic estimator (QE) that is designed to mitigate the lensing biases due to the kinematic and thermal SZ effects. The idea behind the approach is to use inpainting technique to eliminate the cluster's own emission from the large-scale CMB gradient map. In this inpainted gradient map, we fill the pixel values at the cluster location based on the information from surrounding regions using a constrained Gaussian realization. We show that the noise induced due to inpainting process is small compared to other noise sources for upcoming surveys and has negligible impact on the final lensing signal-to-noise. Without any foreground cleaning, we find a stacked mass uncertainty of 6.5% for the CMB-S4 experiment on a cluster sample containing 5000 clusters with $M_{200c} = 2 \times 10^{14}$ $M_\odot$ at $z = 0.7$. In addition to the SZ-induced lensing biases, we also quantify the low mass bias arising due to the contamination of the CMB gradient by the cluster convergence. For the fiducial cluster sample considered in this work, we find that this bias is negligible compared to the statistical uncertainties for both the standard and the modified QE even when modes up to ~ 2700 are used for the gradient estimation. With more gradient modes, we demonstrate that the sensitivity can be increased by 14% compared to the fiducial result quoted above using gradient modes up to 2000.

Keywords: cosmic background radiation – gravitational lensing:weak – galaxies: clusters: general

Contents

1 Introduction **1**

2 Methods **3**
2.1 Lensing estimator 3
2.1.1 The quadratic estimator for CMB-cluster lensing 3
2.1.2 Modification to the QE 5
2.2 Simulation set-up 6

3 Results **7**
3.1 Results from idealized simulations 8
3.1.1 Baseline setup with $\ell_G = 2000$ 8
3.1.2 Effect of cluster convergence bias 9
3.1.3 Boosting lensing S/N with more gradient modes 11
3.2 Effect of cluster correlated foregrounds 11
3.2.1 Mock kinematic SZ signal 13
3.2.2 Kinematic SZ from Sehgal simulations 14
3.2.3 Thermal SZ from Sehgal simulations 14
3.2.4 Discussion and applicability to real data 15

4 Conclusion **16**

1 Introduction

Galaxy clusters are remarkable cosmological probes, with the caveat that their masses must be accurately measured. In general, cluster masses are not directly measured but estimated from mass-observable scaling relations with observed quantities such as the Sunyaev-Zel'dovich(SZ) fluxes for microwave surveys [1, 2]; galaxy richnesses and galaxy velocities for optical surveys [3–5]; or temperatures of the intra-cluster medium and luminosities for X-ray [6, 7] measurements. The mass precision of these methods has been steadily progressing although systematic uncertainties that dominate the cluster mass budget remain around that complex astrophysics that relate the survey observables to the underlying dark matter. Once the mass is correctly measured, the abundance of galaxy clusters as a function of redshift and mass can reveal a great deal of information about the cosmological parameters that control the geometry and growth of structures in the Universe [for a review, see 8].

Weak gravitational lensing can provide an unbiased probe of the matter distribution in a galaxy cluster and, arguably yields the best mass estimate. Consequently, several efforts have been made in the last decade to calibrate the mass-observable scaling relations mentioned above using lensing measurements [recently, 9–11]. Weak-lensing measurements can be performed using two background light sources: galaxies or the cosmic microwave background (CMB). While galaxy weak-lensing measurements provide excellent mass constraints for clusters below $z \leq 1$, CMB-cluster lensing is more effective for distant galaxy clusters. Galaxy weak lensing struggles at these high redshifts due to the difficulty in imaging enough background (i.e. even higher redshift) galaxies at sufficient signal-to-noise (S/N). The CMB, in contrast, originates at $z = 1100$ and backlights all galaxy clusters at any redshift. Thus,

CMB-cluster lensing is expected to yield the best mass measurements for the thousands of distant galaxy clusters [12] expected to be detected by next generation CMB surveys [13, 14]. Unfortunately though, the lensing S/N for a single cluster is low and one must rely on the stacked mass measurement to achieve a reasonable S/N for CMB-cluster lensing.

Lensing of the CMB by a galaxy cluster produces a unique dipole pattern that is oriented along the direction of the background CMB gradient at the cluster location. Several estimators have been proposed in the literature [15–23] to extract the lensing signal and the first detections were made by the Atacama Cosmology Telescope (ACT) [24], South Pole Telescope (SPT) [25], and *Planck* [26] experiments. The measurements have flourished since then and have been used to: constrain the hydrostatic mass bias of clusters [26–28]; constrain the optical richness-mass ($\lambda - M$) relation of galaxy clusters [29–31]; and estimate masses of halos hosting high redshift quasars [32].

Lensing measurements using CMB temperature data, however, are highly susceptible to foregrounds, the cluster Sunyaev-Zel'dovich (SZ) signals in particular [12], and previous works have reduced the SZ-biases at the cost of the lensing S/N [25, 30]. In the standard quadratic estimator (sQE), the lensing reconstruction is performed using two maps: one to measure the large-scale CMB gradient and the other map to measure the small-scale anisotropies [20]. This estimator leverages the fact that gravitational lensing correlates different angular scales that are uncorrelated in the primordial CMB anisotropy. The real world scenario is complicated by the cluster's own emission as such foreground signals are also correlated across different angular scales. But given that the lensing signal is oriented along the direction of the background CMB gradient while the foregrounds have no such dependence, it must be possible to decouple the foregrounds from biasing the lensing results.

The most important source of bias is the cluster thermal SZ (tSZ) signal. The tSZ effect is caused because of the interaction between CMB photons and the hot electrons inside the clusters and has a unique frequency dependance [33]. Recently, Madhavacheril and Hill [34] and Raghunathan *et al.* [31, hereafter R18] used this frequency dependance and modified the quadratic estimator to estimate the background CMB gradient using a tSZ-free map. This tSZ-free gradient QE eliminates the tSZ-induced correlation between the two maps, and thus removes the tSZ bias.

There is also a second SZ effect due to the doppler shifting of CMB photons by the peculiar velocity of the cluster [35], known as the kinematic SZ (kSZ) effect. While smaller on average than the tSZ effect, the kSZ signal has been observed by several experiments [36–43]. Unfortunately, the kSZ signal can not be removed in the same fashion as the tSZ effect because the kSZ effect has the same blackbody spectrum as the CMB. As a result, the tSZ-free gradient QE mentioned above still suffers from a bias due to the kSZ effect. The level of kSZ bias is much smaller than the statistical errors for current measurements [31] but is expected to cause problems for cluster lensing measurements with future low noise CMB datasets [12].

Other cluster emission can also bias the CMB-cluster lensing mass measurements. After the SZ effects, the most significant of these are the thermal dust and synchrotron emission from the cluster's member galaxies. Like the tSZ effect, one might depend on a linear combination of observing frequencies to remove these signals – at the cost of reducing the S/N.

Since the SZ effect and other foregrounds are largely unpolarized [44, 45], working with polarization based estimators is another effective way to handle these foreground biases. Polarization estimators have the advantage of foreground simplicity, but are handicapped by the fact that, for CMB experiments over the next decade, the temperature measurements

have much higher potential S/N. An alternative approach is to combine the temperature and polarization measurements by simply replacing the large-scale CMB gradient map with the polarization E/B-mode maps. However, this approach does not yield the same S/N as temperature-only measurements as the correlation between the primordial CMB temperature and polarization fields is $\lesssim 20\%$.

In this work, we design a modified QE that is robust against bias from all types of cluster emission. We focus on performance with respect to the kSZ and tSZ effect biases, but the scheme should also mitigate dust and radio emission. The idea is to use an inpainted map for the large-scale CMB gradient estimation. In this approach, we replace the pixel values of the CMB gradient map near the cluster location using information from the surrounding regions. We show that the estimator yields unbiased lensing mass estimates on simulations. While, in principle, the inpainting process can introduce additional variance, we show that is minimal and has negligible impact on the lensing S/N compared to the sQE. Inpainting eliminates the spatially-localized foregrounds and distortions, while maintaining the harmonic structure of background CMB. With no cluster SZ signals in the inpainted gradient map, there will no longer be foreground-induced correlations between the two maps used in the QE and as a result no bias in the lensing reconstruction. We also study the bias due to SZ signals beyond the typical cluster size chosen for inpainting that originate from the haloes correlated to the cluster and find them to be negligible. The injected SZ foregrounds for these tests are extracted from N-body simulations [46] to capture all real-world effects. Finally, an added advantage of inpainting is that it also allows us to estimate the unlensed CMB gradient without being influenced by the cluster's own convergence field. Hu *et al.* [20] noted that the magnification created by cluster lensing led to an underestimation of the background CMB gradient at the cluster location. We quantify this mass bias in the sQE and compare it to the inpainted gradient estimator.

The paper is organized as follows. The standard quadratic lensing estimator, its modifications, and a brief description of the inpainting technique are described in §2.1. This is followed by an outline of simulations used to validate the cluster lensing pipeline in §2.2. In §3.1, we compare our estimator to the sQE and tSZ-free gradient estimators, and then quantify the effect of cluster convergence bias. The tests to check robustness of the estimator to foreground induced lensing biases are in §3.2. We discuss the potential of the estimator with future CMB datasets and conclude in §4. Throughout this work adopt the ΛCDM cosmology obtained from *Planck* chain that combines the *Planck* 2015 temperature, polarization, lensing power spectra with BAO, H0, and SNe data (TT,TE,EE+lowP+lensing+ext in Table 4 of Planck Collaboration *et al.* [47]).

2 Methods

We introduce the standard lensing quadratic estimator and then present the inpainting technique. This is followed by an overview of the simulations used to validate the lensing pipeline.

2.1 Lensing estimator

2.1.1 The quadratic estimator for CMB-cluster lensing

Gravitational lensing of the CMB remaps the unlensed CMB temperature or polarization fields $\tilde{X}(\hat{\mathbf{n}})$ ($X \in [T, E, B]$) to new positions based on the deflection angle $\alpha(\hat{\mathbf{n}}) = \nabla\phi$ corresponding to the underlying lensing potential ϕ. The total power is conserved by gravitational

lensing. Thus, the observed CMB field X at a given position can be written as [48]

$$X(\hat{\mathbf{n}}) = \tilde{X}[\hat{\mathbf{n}} + \nabla\phi(\hat{\mathbf{n}}))] \sim \tilde{X}(\hat{\mathbf{n}}) + \nabla\tilde{X}(\hat{\mathbf{n}}) \cdot \nabla\phi(\hat{\mathbf{n}}), \tag{2.1}$$

where we have dropped higher order terms in the Taylor expansion to obtain the expression on the right.

It is convenient to transform the CMB and deflection angle fields to Fourier space, $\tilde{X}_\ell$ and α_ℓ, where we have used the subscript ℓ to indicate the Fourier transformations. In Fourier space, lensing convolves the unlensed CMB field $\tilde{X}_\ell$ and the underlying deflection angle field α_ℓ. Thus, lensing correlates the otherwise independent modes. The basic idea behind the QE is to use these lensing induced correlations to reconstruct the underlying lensing convergence signal $\kappa = -\nabla^2\phi$.

The QE looks at the product of two maps, G and L, to reconstruct the lensing signal [20]:

$$\hat{\kappa}_{\boldsymbol{\ell}} = -A_\ell \int d^2\hat{\mathbf{n}}\; e^{-i\hat{\mathbf{n}}\cdot\boldsymbol{\ell}}\; \mathrm{Re}\left\{\nabla \cdot [G(\hat{\mathbf{n}})L^*(\hat{\mathbf{n}})]\right\}. \tag{2.2}$$

The two maps, which would be independent in the absence of lensing, are both derived from the CMB temperature map T according to:

$$G_{\boldsymbol{\ell}} = i\boldsymbol{\ell}\, W_\ell^G\, T_{\boldsymbol{\ell}}^G, \tag{2.3}$$

$$L_{\boldsymbol{\ell}} = W_\ell^L\, T_{\boldsymbol{\ell}}^L. \tag{2.4}$$

Here the first map $G_{\boldsymbol{\ell}}$ corresponds to a large-scale gradient map while $L_{\boldsymbol{\ell}}$ is a CMB temperature map weighted optimally to extract the small-scale lensing signal. The notational choice here to distinguish T^G and T^L prefigures the modified versions of the QE; in the original estimator the same temperature map is used for both legs, $T \equiv T^G = T^L$. The optimal linear filters W_ℓ^G and W_ℓ^L are designed to maximize the lensing S/N [20]:

$$W_\ell^G = \begin{cases} C_\ell^{\rm unl}(C_\ell + N_\ell^G)^{-1} & , \ \ell \le \ell_{\rm G} \\ 0 & , \ \text{otherwise} \end{cases}$$

$$W_\ell^L = (C_\ell + N_\ell^L)^{-1}. \tag{2.5}$$

In the above equation, $(C_\ell^{\rm unl})C_\ell$ represent the (un)lensed CMB temperature power spectra. In this work, we calculate these spectra using the Code for Anisotropies in the Microwave Background (`CAMB`[1]) software [49]. N_ℓ is the beam-deconvolved noise power spectrum for a given map.

Note that in Eq. (2.5), the weights for the gradient modes above ℓ_G are set to zero. As pointed out by Hu *et al.* [20], the CMB gradient at the cluster location is under-estimated due to the magnification of the CMB by the underlying cluster convergence and this leads to bias in the reconstructed cluster mass. Henceforth, we will refer to this effect as the *cluster convergence bias*. Since scales larger than a typical cluster scale (few arcminutes) are relatively unaffected by the convergence signal, the bias can be mitigated by simply removing the more contaminated small-scale modes in the gradient estimation. Subsequently, several works have set $\ell_G = 2000$ [23, 24, 31] in their analyses following Hu *et al.* [20]. This heuristic choice of $\ell_G = 2000$ comes from the fact that the small-scale power of the unlensed CMB is exponentially damped [50] and excluding gradient modes $\ell \ge 2000$ does not significantly

[1] https://camb.info/

degrade the lensing S/N [also see Fig. 1 of 20]. We return to this discussion in § 3.1.3 and demonstrate that inpainting allows more modes to be included (and thus adds sensitivity) without being affected by the cluster convergence bias.

2.1.2 Modification to the QE

As mentioned earlier, the bias due to the cluster tSZ signal can be mitigated by estimating the background CMB gradient from a tSZ-free CMB map. For example, R18 combined the 95 and 150 GHz channels from the SPTpol experiment to clean the tSZ signal from the gradient map. While the tSZ-free gradient approach works, in principle, to eliminate the bias due to cluster tSZ signal, it is not suitable to remove the lensing induced bias due to cluster kSZ signal as, unlike tSZ, kSZ has no average frequency dependence.

In this work, we follow a different approach. We replace the tSZ-free gradient map with a inpainted gradient map in which we fill the pixel values at the cluster location with the information from the surrounding region. This inpainted gradient approach has the following two key merits compared to a tSZ-free gradient. First, the approach can help in mitigating the kSZ-induced lensing bias; and second, the noise and other frequency dependent foregrounds in the gradient map are not enhanced.

Based on Benoit-Lévy *et al.* [51] we define two regions at distances $R \leq R_1$ and $R_1 < R \leq R_2$ from the cluster center. We fill the CMB temperature values T_1 in region R_1 based on the temperature T_2 in R_2 using a constrained Gaussian realization as

$$\hat{T}_1 = \tilde{T}_1 + \hat{\mathbf{C}}_{12}\hat{\mathbf{C}}_{22}^{-1}(T_2 - \tilde{T}_2) \tag{2.6}$$

where $\tilde{T}_1, \tilde{T}_2$ are the temperature values in the two regions from a random Gaussian realization of the CMB and we use hats to represent the estimated quantities. The above operation is possible since the primordial CMB exhibits coherence over degree scales and the covariance matrix $\hat{\mathbf{C}}_{XY}$ of the CMB fields between two regions X, Y can be calculated using simulations as

$$\hat{\mathbf{C}}_{XY} = \hat{\mathbf{C}}_{YX} = \frac{1}{n-1}\sum_{i=0}^{n}(\mathbf{G}_i - \langle\mathbf{G}\rangle)(\mathbf{G}_i - \langle\mathbf{G}\rangle)^T \tag{2.7}$$

where the index i runs over the number of simulations and $\mathbf{G} = \begin{pmatrix}\mathbf{G}_X \\ \mathbf{G}_Y\end{pmatrix}$ is the concatenated vector of a Gaussian realization of the large-scale structure (LSS) lensed CMB map in the two regions. The details about our simulations can be found in Raghunathan *et al.* [12] and are discussed briefly in the following section § 2.2.

The effectiveness of the estimator depends on how close the inpainted values $\hat{T}_1$ are to the true background CMB gradient T_1. Of course, inpainting introduces extra noise and, as expected, the difference between true and inpainted values is higher when the input map is noisier. In addition, the process of inpainting introduces mode mixing and, as a result, even though we are only interested in reconstructing the large-scale gradient modes, it is not trivial to avoid the contamination from the noisy small-scale modes. Consequently, we low-pass filter (LPF) the input CMB map T to remove modes $\ell > \ell_{\rm LPF}$ before performing the inpainting process.

We note that the size of the inpainting radius $R1$, in principle, depends on the angular size of the cluster and must be chosen dynamically based on the cluster mass and redshift in case of real data. Here, we conservatively set $R1 = 4'$ to encompass the clusters from a

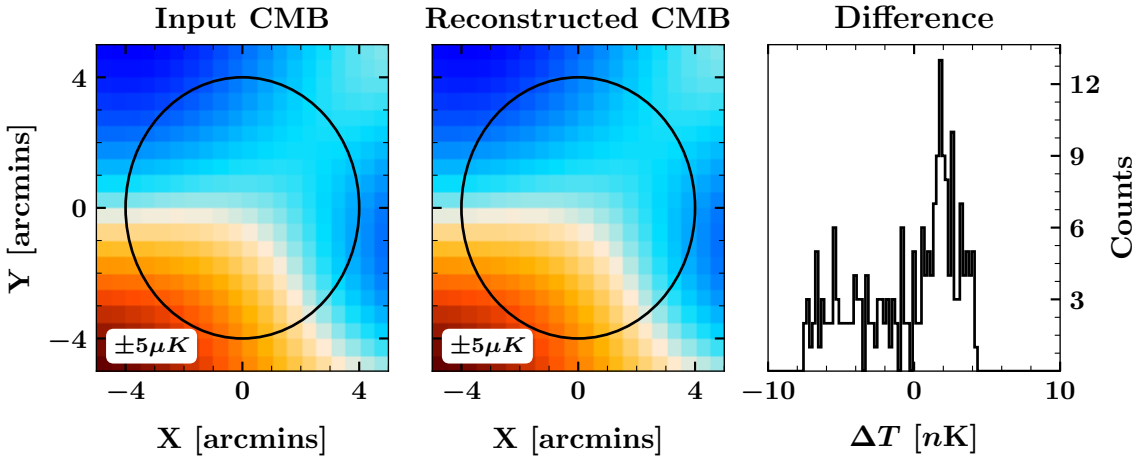


Figure 1. Illustration of the reconstructed CMB temperature using the inpainting technique. The first two panels are $10' \times 10'$ cutouts of the input and the reconstructed signals. They have been LPF at $\ell_{\rm LPF,R1}$ and include signals from CMB, foregrounds, and noise. The black circle is the ($R1 = 4'$) inpainted region. The right panel shows the pixel differences between the two within the inpainted region in nano-Kelvin units and is much smaller than the noise in the input CMB map in the left panel.

wide range of masses and redshifts. As an example, for the clusters considered in the current work $R1 \sim \theta_{200c}$ where $\theta_{200c} = R_{200c}/D_A(z)$ with D_A being the angular diameter distance to the cluster redshift z. For massive nearby clusters, however, $R1$ must be higher as we demonstrate in later sections. For the LPF, we use $\ell_{\rm LPF} \equiv \ell_{\rm LPF,R1} = 2700$ which corresponds to the size of the inpainting radius $R1$. The choice of $\ell_{\rm LPF,R1}$ is well motivated without much S/N penalty given that: (a) no information can be recovered from the inpainted map below the scale corresponding to $R1$; and (b) the small-scale primordial CMB is exponentially damped [50]. The only drawback of the LPF is the undesired ringing effects in the presence of a bright SZ signal at the cluster location. In §3.2, we describe additional filtering steps in the input CMB map T to attenuate these filtering artifacts and cleanly reconstruct the cluster convergence. The outer region $R2$ is set to $30'$ and we perform 50,000 simulations to estimate the covariance matrix. We have confirmed that increasing R2 has negligible impact on our results. An illustration of the inpainting process is given in Fig. 1 where the first two panels are the $10' \times 10'$ cutouts from the original and reconstructed CMB maps around the desired region chosen for inpainting. Both the panels include signals from CMB, and variance from noise/foregrounds and have been LPF at $\ell_{\rm LPF,R1}$. The histogram in the right is the difference between the two panels within $R1$ (black circle). The residual shown in nano-Kelvin units is negligible compared to the noise in the input CMB map.

2.2 Simulation set-up

Following the previous work [12], we create Gaussian realizations of the CMB temperature map using the LSS lensed power spectra C_ℓ^{TT} generated using `CAMB` [49] software for the fiducial *Planck* 2015 cosmology. In these simulations, we neglect the small non-Gaussianities arising due to the lensing from intervening LSS. The simulations are done in the flat-sky approximation and are $120' \times 120'$ wide with a pixel resolution of $0.'5$. The simulations

are then lensed using galaxy cluster convergence profiles described below. We add Gaussian realizations of the foregrounds not associated to the galaxy cluster using the measurements made by the SPT-SZ experiment [52]. The foregrounds include dusty and radio galaxy emissions, and the SZ signals from the unresolved haloes in the SPT-SZ 90 and 150 GHz channels. In §3.2, we also include the cluster SZ signals to quantify the level of bias induced in the reconstructed lensing signal. In all cases, the foregrounds remain unlensed by the galaxy cluster. Next, we convolve the simulated maps with a Gaussian beam with $\theta_{\rm FWHM} = 1'.4$ and add instrumental noise of $\Delta_T = 1.5\mu{\rm K}'$ matching the expected configuration[2] of the wide area survey of the CMB-S4 [13] experiment. For the tSZ-free maps, we set the noise to a slightly higher level $\Delta_T = 4.6\mu{\rm K}'$ and use $\theta_{\rm FWHM} = 2'$. This higher noise level is chosen to match the CMB-S4 tSZ-free map made from the combination of 90 and 150 GHz channels. Adding data from other available channels will tend to reduce the noise level slightly.

To model the galaxy cluster profile, we use the Navarro-Frenk-White (NFW) [53] formalism given in Eq. (2.8)

$$\rho\,(r) \,=\, \frac{\rho_0}{\left(\frac{r}{R_{\rm s}}\right)\,\left(1+\frac{r}{R_{\rm s}}\right)^2}, \tag{2.8}$$

where $R_{\rm s}$ is the scale radius, ρ_0 is the central cluster density, and $c_{200}(M,z) = R_{200}/R_{\rm s}$ is the halo concentration parameter. To obtain the NFW convergence profile $\kappa(M,z)$ for the NFW halo, we simply plug in the analytic expression given in Eq. (2.8) of Bartelmann [54]. We quote the cluster masses with respect to the radius R_{200c}, which encompasses the region within which the average mass density is 200 times the critical density of the Universe at the cluster redshift z. For simplicity, we fix the cluster masses and redshifts to: $M_{200c} = 2\times10^{14}$ M$_\odot$ and $z = 0.7$ with $c_{200} = 3.12$ calculated using Duffy *et al.* [55] formula.

3 Results

In § 3.1, we compare the inpainted gradient to the sQE and tSZ-free gradient QE followed by a discussion about approaches to improve lensing S/N. Next, in § 3.2, we quantify the impact of the cluster SZ signals on the reconstructed lensing maps and how they can be mitigated using the inpainted gradient estimator.

Throughout this section, we quote the mass constraints obtained by stacking the reconstructed lensing signal $\hat{\boldsymbol{\kappa}}$ from 5000 galaxy clusters. The lensing masses quoted in the subsequent figures are the median values obtained from ten such simulations, each with 5000 stacked clusters. We bin the stacked signal into 10 linearly-spaced radial bins $\hat{\kappa}(\theta)$ with $\Delta\theta = 1'$. We compare the radially-binned $\hat{\kappa}(\theta)$ to the fiducial NFW convergence $\kappa(M,z)$ profiles binned radially into $\kappa(\theta)$ and obtain cluster mass constraints using

$$-\,2\ln\mathcal{L}(\hat{\kappa}(\theta)|M) = \sum_{\theta=0}^{10'} \left[\hat{\kappa}(\theta)-\kappa(\theta)\right]\hat{\mathbf{C}}^{-1}\left[\hat{\kappa}(\theta)-\kappa(\theta)\right]^T \tag{3.1}$$

To obtain the covariance matrix we use a jackknife re-sampling technique

$$\hat{\mathbf{C}} = \frac{N-1}{N}\sum_{j=1}^{N=1000}\left[\hat{\kappa}_j(\theta)-\langle\hat{\kappa}(\theta)\rangle\right]\left[\hat{\kappa}_j(\theta)-\langle\hat{\kappa}(\theta)\rangle\right]^T, \tag{3.2}$$

[2]We use the projected performance of the CMB-S4 150 GHz channel for the Large-area Survey 03 specified in the following link as on 13 March 2019: https://cmb-s4.org/wiki/index.php/Survey_Performance_Expectations#Large-area_Survey_Performance_Expectation_03

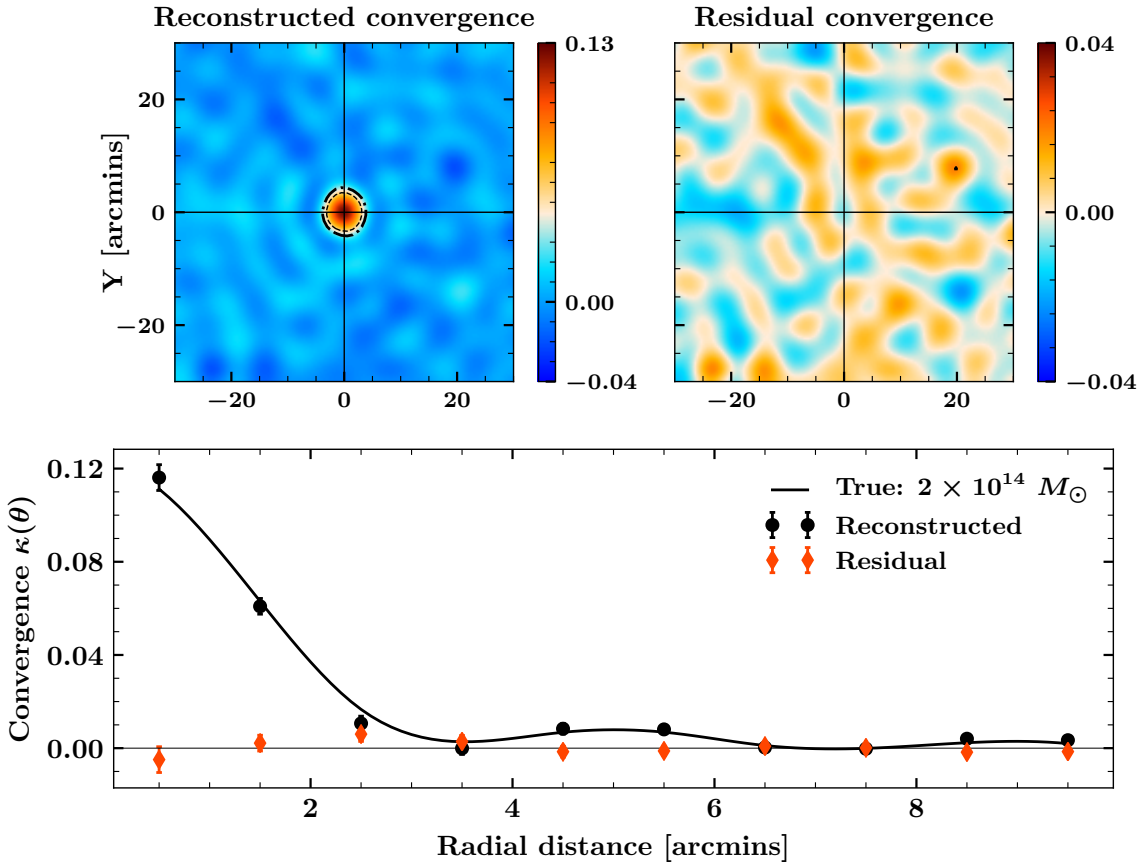


Figure 2. Top left panel shows the reconstructed stacked convergence maps of 5000 clusters using the inpainted gradient estimator. The difference between the true cluster convergence and the left panel is shown in the top right panel. The radially-binned reconstructed (black circles), true (black solid curve), and residual (orange diamonds) convergence profiles are in the bottom panel. The residual signal is consistent with a null signal, probability to exceed (PTE) = 0.55, implying that the reconstructed convergence signal is unbiased. The contour lines correspond to regions that are $\geq 3, 5\sigma$ from the background.

where $\hat{\kappa}_j(\theta)$ is the azimuthally binned stacked convergence of all the clusters in the j^{th} sub-sample and $\langle\hat{\kappa}(\theta)\rangle$ is the ensemble average of all the 1000 sub-samples. For parameter constraints, we sample the likelihood space and report the median mass values and 1σ uncertainties from the 16^{th} and 84^{th} percentiles.

3.1 Results from idealized simulations

First, we discuss the mass constraints from all the three QE for idealized simulations with no cluster foreground signals and $\ell_G = 2000$. Next, we modify ℓ_G to include more gradient modes and quantify the cluster convergence bias and the improvement in lensing S/N.

3.1.1 Baseline setup with $\ell_G = 2000$

For illustrative purposes, we show the reconstructed convergence map using the inpainted gradient QE in Fig. 2. The stacked convergence map is in the top left and the residual,

difference between the true cluster convergence and the left panel, on the right. In the bottom panel, we show radial profiles of the true (black curve), reconstructed (black circle), and residual (orange diamond) convergence signals. The PTE value of the residual signal is 0.55 ($\chi^2_{\rm null} = 7.9$ for 9 degrees of freedom), indicating that it is consistent with random fluctuations.

Next, we compare the mass constraints from three estimators in the top panel of Fig. 3: sQE (orange hexagon), tSZ-free gradient (red diamond), and inpainted gradient (black circle). All the three estimators return a cluster mass that is $\leq 0.1\sigma$ from the input mass shown as black dash-dotted line. With the assumed experimental configuration and with no foreground cleaning for the 150 GHz channel, we note that the uncertainty in the stacked mass of our cluster sample is roughly $\Delta M/M \sim 6.5\%$. The fact that the sQE and inpainted gradient QE return similar constraints indicates that the noise from the inpainting process is negligible. Another subtle thing to note is that the tSZ-free gradient, created by combining multiple frequency channels has a higher noise ($\Delta_T = 4.'5\mu$K$'$) compared to the fiducial 150 GHz maps ($\Delta_T = 1.'4\mu$K$'$). This could result in a reduced S/N for the tSZ-free gradient QE compared to the others. At first glance, such a trend is not evident when restricting the gradient modes to $\ell_G = 2000$. We will come back to this discussion in following sections where we analyze the changes in the S/N upon modifying ℓ_G.

3.1.2 Effect of cluster convergence bias

As mentioned in previous sections, the gradient estimation is limited to scales larger than $\ell \leq \ell_G$ for QE in general. This LPF is introduced to mitigate the cluster convergence bias arising from an underestimated gradient at the cluster location at the cost of a loss in sensitivity. In the baseline case in §3.1.1, we adopt $\ell_G = 2000$ which is reasonable for less massive clusters [20]. But for massive low-redshift clusters which span a larger angular extent on the sky, $\ell_G = 2000$ may not be sufficient and can result in a mass bias of the order of few per cent [20]. The goal of this section is to study the effect of the cluster convergence bias for different cluster samples and understand if the sensitivity can be improved without being affected by the bias.

We start this test by investigating the recovered lensing mass from all three estimators for various values of ℓ_G. This is shown in the bottom panel of Fig. 3. For the fiducial setup, the effect of the magnification bias on modes $\ell < \ell_G (= 2000)$ is negligible. However, when the value of ℓ_G is increased (2000 through 5000), the recovered mass systematically shifts to lower values for both sQE (hexagon) and tSZ-free gradient (diamond) points, which could be a hint of the cluster convergence bias. The inpainted gradient estimator (open circles), which is devoid of cluster convergence, is not affected and returns an unbiased mass for all values of ℓ_G. While we impose $\ell_G > \ell_{\rm LPF,R1}$, we note that the size of the inpainting radius $R1$ acts as an ultimate LPF for the inpainted gradient QE and no useful gradient information can be recovered from scales smaller than $\ell_{\rm LPF,R1} \geq 2700$.

Results for more massive cluster samples: Since the level of bias for sQE and tSZ-free gradient QE is much smaller than the statistical errors, the results are inconclusive for our fiducial cluster sample. So we extend our analysis to more massive clusters to exaggerate the cluster convergence bias, and measure it in a computationally reasonable sample size. For this test, we use samples of clusters that are 4×10^{14} $M_\odot$ (i.e. twice as massive as our fiducial sample) at redshifts $z = 0.7$ and $z = 1.5$. The results are presented in Tab. 1. We chose two different radii $R1 = 4'$ and $5'$ for the inpainted gradient estimator and, to make a fair comparison, set $\ell_G = 2700$ and 2160 corresponding to the scale of $R1$ in the sQE. Our results

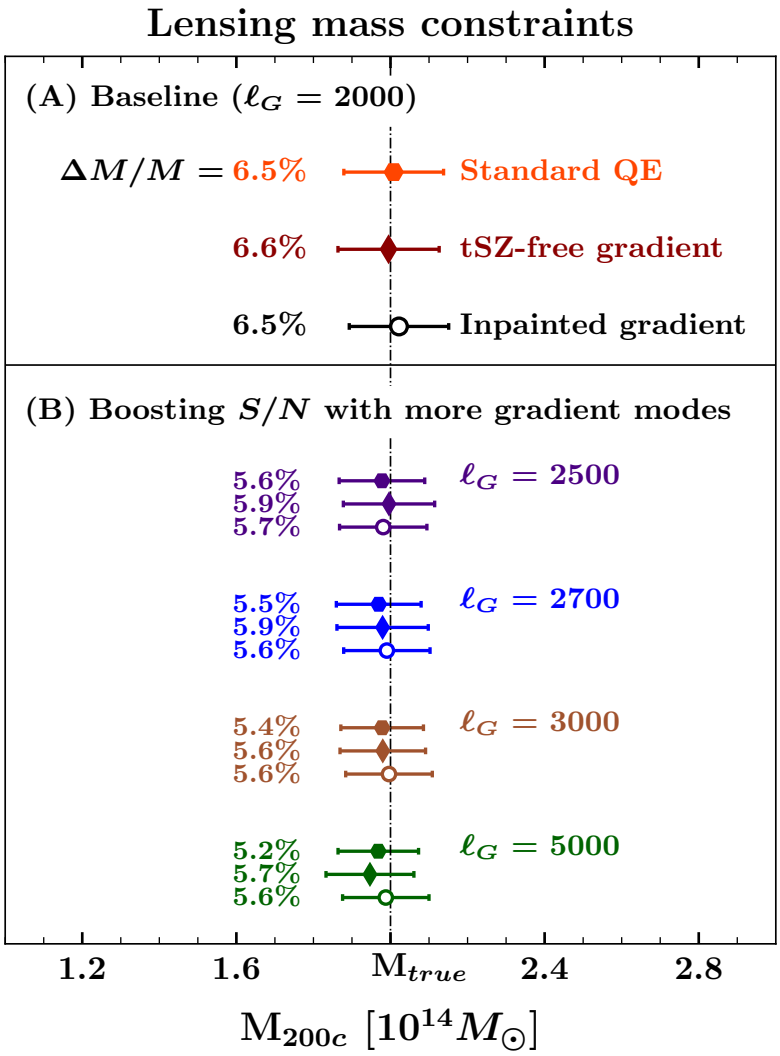


Figure 3. *Top panel:* The mass constraints obtained from the stacked convergence signal for the baseline case (ℓ_G = 2000) with the sQE (hexagon), tSZ-free gradient (diamond), and inpainted gradient (circle) QE estimators. *Bottom panel:* The improvement in mass precision by including more modes for the gradient estimation. For the inpainted gradient QE, the S/N saturates after ℓ_G = 2700, which roughly corresponds to the scale of the inpainting radius $R1 = 4'$.

suggest that the systematic bias due to cluster convergence is smaller than the statistical uncertainty for both the sQE and the inpainted gradient when ℓ_G is set to 2700 or smaller. When $\ell_G \geq 3000$ in sQE, the bias increases and becomes comparable to the statistical errors. These results are consistent with the observations from Fig. 3 for the fiducial cluster sample. For even massive clusters with $M_{200c} = 8 \times 10^{14}$ M$_\odot$ at $z = 0.7$, the bias is evident. The lensing inferred mass is shifted low by 7% for both the sQE ($\ell_G = 2700$) and inpainted gradient ($R1 = 4'$) estimators. The bias vanishes when $R1$ is set to $8'$ for inpainted gradient or equivalently $\ell_G = 1350$ for sQE with a significant ($\times$ 1.75) hit in S/N.

Based on the above results we conclude that it is safe to set $\ell_G \sim 2700$ when trying to reconstruct the convergence signal, of clusters with $M_{200c} \lessapprox 4 \times 10^{14}$ M$_\odot$, using next-generation

Table 1. Bias in the reconstructed lensing masses due to the cluster convergence signal. We consider two clusters samples with $M_{200c} = 4 \times 10^{14}$ $M_\odot$ at redshifts $z = 0.7$ and $z = 1.5$.

Estimator	ℓ_G	$z = 0.7$		$z = 1.5$	
		Lensing mass [10^{14} $M_\odot$]	Bias % (frac. of σ)	Lensing mass [10^{14} $M_\odot$]	Bias % (frac. of σ)
Inpainted gradient QE	2700 ($R1 = 4'$)	3.99 ± 0.14	0.2% (0.1σ)	3.96 ± 0.13	1.0% (0.3σ)
	2160 ($R1 = 5'$)	4.01 ± 0.15	0.2% (0.1σ)	3.99 ± 0.14	0.2% (0.1σ)
Standard QE	3000	3.93 ± 0.13	1.8% (0.5σ)	3.81 ± 0.12	4.8% (1.5σ)
	2700	4.04 ± 0.13	1.1% (0.3σ)	3.93 ± 0.12	1.8% (0.6σ)
	2160	4.05 ± 0.15	1.2% (0.3σ)	4.03 ± 0.14	1.1% (0.2σ)

datasets without worrying about convergence bias. The inpainted gradient QE, however, does not have specific advantage over the sQE with respect to the cluster convergence bias. The gain in the lensing S/N with a higher ℓ_G is discussed in the next section.

3.1.3 Boosting lensing S/N with more gradient modes

Now that we have an estimate of the cluster convergence bias, we could improve the S/N of the estimated convergence by including more modes in the gradient estimation. This improvement is shown in the bottom panel of Fig. 3, where we show the mass constraints for different values of ℓ_G from all the three QE. From the figure, it is evident that the S/N increases for all estimators when more gradient modes are included. The rate at which the S/N improves and the saturation point, however, differs.

When ℓ_G is set to 2500 (purple), we obtain 15% better mass constraints for sQE, 9% for tSZ-free gradient, and 14% for inpainted gradient compared to the fiducial cases in the top panel with $\ell_G = 2000$. The reason for a significantly smaller improvement in tSZ-free gradient QE is attributed to the higher noise in tSZ-free map compared to the 150 GHz or the inpainted gradient map. The rate at which the S/N improves is only gradual for even higher ℓ_G due to the diffusion damping of the small-scale CMB [50] (also see Fig. 1 of Hu *et al.* [20]).

There is no improvement in S/N for the inpainted gradient when $\ell_G \geq 2700$, emphasising no useful gradient information can be recovered from scales smaller than the size of the inpainting radius $R1 = 4'$. As expected, the S/N saturates at a higher ℓ_G for both sQE and tSZ-free gradient estimators. Since the tSZ-free gradient map is noisier than the 150 GHz map, the tSZ-free gradient QE attains saturation earlier than the sQE.

3.2 Effect of cluster correlated foregrounds

Now we turn to the effect of cluster-correlated foregrounds: kSZ and tSZ signals from the galaxy cluster on the standard, tSZ-free gradient, and the inpainted gradient QE. As discussed before, the cluster foregrounds introduce a bias in the lensing reconstruction using QE because of the correlations between foregrounds present in the maps in the two legs of the estimator. As a result, we can expect both the cluster tSZ and kSZ to bias the standard QE. The tSZ-free gradient QE should only be sensitive to the bias from cluster kSZ which cannot be cleaned by combining data from multiple frequencies. Cluster tSZ should have no effect in this case since, by construction, tSZ-free gradient QE is free from tSZ in the first leg of the estimator. As we show below, the inpainted gradient QE is insensitive to both cluster tSZ and kSZ,

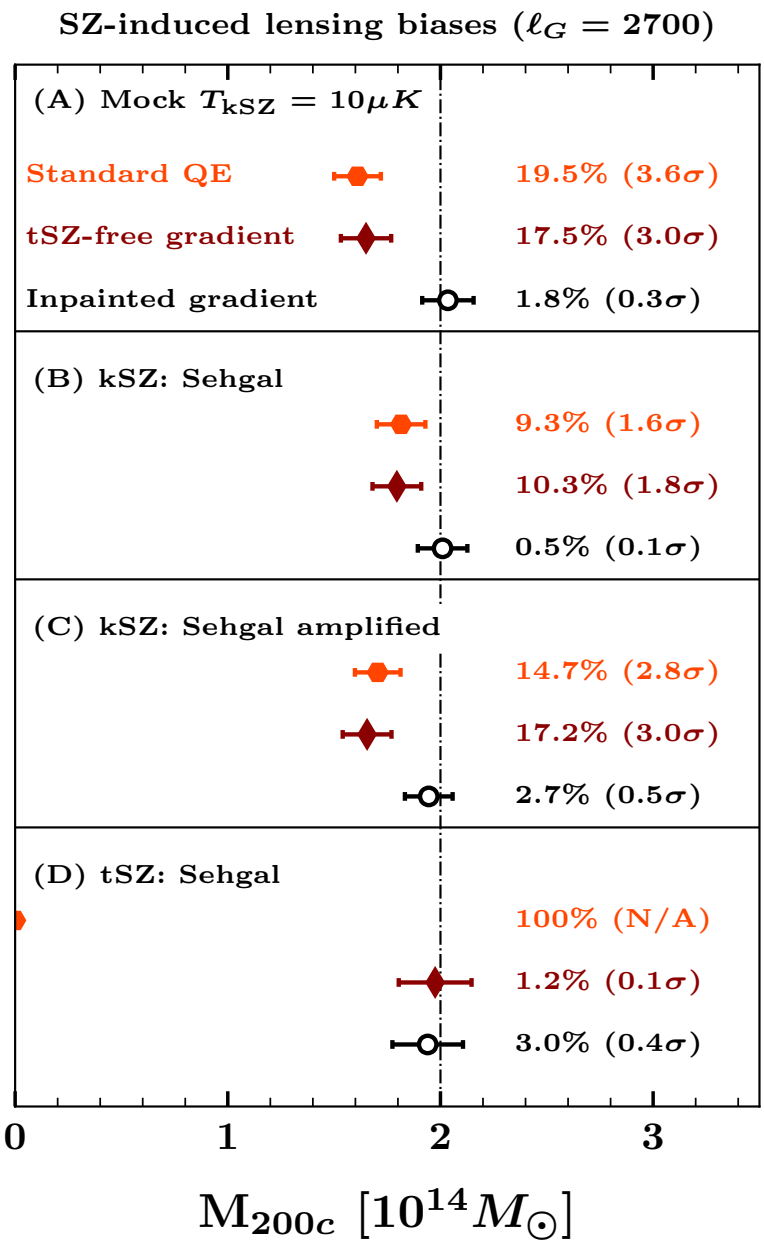


Figure 4. Same as Fig. 3, but after including cluster SZ signals in the CMB maps. The data points for the standard, tSZ-free gradient, and inpainted gradient are shown as orange hexagons, red diamonds, and black circles. The top three panels are for varying level of kSZ signals: top panel is for a mock kSZ with $T_{\rm kSZ} = 10\mu K$; second panel is using kSZ from S10 simulations; third panel also uses S10 simulations but after amplifying them using the signals corresponding to clusters with mass 4×10^{14} $M_\odot$. In all cases, the inpainted gradient provides an unbiased reconstruction of the cluster mass while the bias in the sQE and tSZ-free gradient grow with the amplitude of the kSZ signal. The bottom panel is for tSZ from S10 simulations. The sQE is heavily biased due to cluster tSZ while the tSZ-free gradient and inpainted gradient are unbiased.

as they are removed using the inpainting technique in the first leg. We use $\ell_G = 2700$ for all the tests in this section but also show the reduction in foreground bias for the sQE with $\ell_G = 2000$.

First, we inject simple mock kSZ signals where we simply alter the pixel temperatures by a certain amount. This is followed by realistic foreground signals from Sehgal *et al.* [46, hereafter S10] simulations. To this end, we extract half-arcminute-resolution $120' \times 120'$ cutouts from the 90 and 150 GHz kSZ and tSZ S10 simulations around halos corresponding to the cluster sample used here. We also check the effect of enhanced foregrounds using SZ signals from more massive clusters in the S10 simulations. We scale the tSZ power down from the S10 simulations by a factor of 1.75 to match the George *et al.* [52] measurements. These foreground cutouts are added to our simulated galaxy-cluster-lensed CMB datasets and processed in the standard manner as discussed in §2.2.

Note that as mentioned in §2.1.2, we LPF the true CMB map at $\ell_{\rm LPF,R1} = 2700$ before the inpainting process to reduce the mode coupling introduced from noisy higher multipoles. This approach works well in the absence of cluster emission as shown in previous sections. When the cluster emissions are included, particularly tSZ, the LPF can introduce ringing, making the inpainting process problematic. There are several known ways of handling this such as, tapering the data with an apodization window or altering the filter to have a smoother roll-off. In this work, we simply remove an estimate of the SZ signal from the gradient map before the LPF step. We explain this in the subsequent sections below. No such approximate foreground removal has been performed for the standard or the tSZ-free gradient QE.

3.2.1 Mock kinematic SZ signal

We start by injecting a mock kSZ signal into the simulated CMB maps which is simply a temperature increment $T = T_{\rm CMB} + T_{\rm kSZ}$ in the pixels that are within $2'$ from the cluster center. The simulated map is then beam-convolved and processed as in the fiducial case. If unaccounted for, the kSZ signal, which is present in both the maps (G and L), will introduce undesired correlations that are otherwise absent. The inpainted gradient approach with $R1 = 4'$ should encompass this mock kSZ signal and must remain insensitive to it.

To ensure that the LPF step before inpainting does not introduce filtering artifacts, we remove an estimated cluster kSZ signal using the aperture photometry (AP) technique [56, 57]. In this technique, an estimate of $\hat{T}_{\rm kSZ}$ is obtained by computing the difference between the integrated temperature values in two concentric shells $T_{\theta_1} - T_{\theta_2}$ around the cluster where we set $\theta_1 = 1.'4$ and $\theta_2 = \sqrt{2}\theta_1$. This kSZ estimate $\hat{T}_{\rm kSZ}$ is removed from the pixels within θ_1 in the gradient map before the LPF step in the inpainting process. We would like to emphasise that our goal here is simply to remove a significant portion of the kSZ signal for an efficient inpainting and not to reproduce the true kSZ profiles. As a result, the AP technique serves our purpose well here.

The results of this test are presented in the top panel of Fig. 4. We use $T_{\rm kSZ} = 10\mu K$. No noticeable bias is detected for the inpainted gradient while the standard and tSZ-free gradient are biased low by more than 17%. The bias reduces to $\sim 6\%$ when a lower $\ell_G = 2000$ is adopted. The low mass bias in the standard and tSZ-free gradient QE grow with the level of $T_{\rm kSZ}$ and is $> 50\%$ for $T_{\rm kSZ} = 20\mu K$. With the inpainted gradient QE, the shift in the reconstructed lensing mass is smaller than the statistical uncertainty ($< 0.4\sigma$) for $T_{\rm kSZ} = 20\mu K$. While it is not realistic that all clusters in the sample have such large velocities and consequently a strong kSZ signal, this test serves to emphasise the robustness of the estimator.

3.2.2 Kinematic SZ from Sehgal simulations

In the previous section, we simply modified the temperature of the pixels within $2'$ from the cluster center to input the kSZ signal into the simulated CMB maps. However, it is plausible that the true kSZ profile of the clusters and the correlated haloes has a much wider spread that can extend beyond the inpainting radius $R1$. If this is true, then the kSZ signal outside $R1$ will not only induce kSZ-bias in the lensing reconstruction but also contaminate the inpainting process. To test this case, we use the kSZ signal corresponding to clusters in our sample from the S10 simulations. As above, we use the AP technique to remove $\hat{T}_{\rm kSZ}$ in the gradient map before inpainting. We also repeated this test with an amplified kSZ signal corresponding to clusters that are twice as massive ($M_{200c} = 4 \times 10^{14}$ M$_\odot$). These two cases are shown in the second and the third panels of Fig. 4.

Like in the case of mock kSZ, the lensing mass from the inpainted gradient remains insensitive to the kSZ signal and the obtained lensing masses are $\leq 0.1\sigma$ and 0.5σ from the true value in the two cases. The lensing mass from the standard and tSZ-free gradient QE are biased low by $\sim 10\%$ in the second panel (3% for $\ell_G = 2000$ instead of 2700) and the bias increases to $\gtrapprox 15\%$ (third panel) when the kSZ signal is amplified.

When using the kSZ signal from even more massive clusters $M_{200c} \geq 5.9 \times 10^{14}$ M$_\odot$, the lensing mass from the inpainted gradient is low by 7.2% (1.3σ). However, the number of haloes above this mass drops to less than 50 in the S10 simulations and our results could be limited by the sample variance. Nevertheless, the bias becomes much smaller than the statistical error when a larger $R1 = 6'$ is adopted for the inpainting process.

3.2.3 Thermal SZ from Sehgal simulations

In the same spirit, we also checked the effect of tSZ signal from the cluster and correlated haloes [58, 59] that extend outside the inpainting radius $R1$ using S10 simulations. As in the kSZ case, we remove a tSZ estimate from the gradient map before the LPF step in the inpainting process. Given that tSZ is much brighter and more extended than kSZ, we do not rely on a simple AP technique to filter tSZ. Instead, motivated by the work of Patil *et al.* [60], we fit a 2D circular Gaussian model to the cluster tSZ signal and subtract that from the CMB map before the LPF step. The Gaussian model has four free parameters: amplitude, x and y centroids, and the width. Note that we perform this filtering in the gradient G map unlike Patil *et al.* [60] in the L map. While a Gaussian template may not represent the true profile of the cluster tSZ signal, the main objective of the template subtraction is to remove a significant amount of tSZ signal to reduce filtering artifacts from the LPF. We perform the fitting with the 150 GHz tSZ map that is devoid of CMB by combining the the 90 and 150 GHz channels.

The tSZ tests are presented in the bottom panel of Fig. 4. The sQE, as expected, is heavily biased in the presence of tSZ. Reducing $\ell_G = 2000$ does not help much in the case of tSZ bias for the sQE. On the other hand, the inpainted gradient estimator after removing a Gaussian tSZ model in the gradient map, works well even in the presence of tSZ. We obtain a lensing mass of $M_{\rm lens} = 1.94 \pm 0.16 \times 10^{14}$ M$_\odot$ consistent with the true cluster mass. If we replace the Gaussian fitting with the AP technique like in the kSZ case, the recovered lensing mass is biased low by 3.5σ for our fiducial inpainting radius $R1 = 4'$. The bias is reduced when $R1$ is increased. As a further test, we amplified the tSZ signal by injecting signals corresponding to clusters with $M_{200c} = 4 \times 10^{14}$ M$_\odot$. The inpainted gradient returns a lensing mass of $M_{\rm lens} = 2.18 \pm 0.29 \times 10^{14}$ M$_\odot$ for this case.

For both the above masses, we note that the mass constraints are weakened (8.2% as opposed to 5.6% in Fig. 3) when Sehgal tSZ haloes are added to the simulations. As observed previously in R18, this extra variance is due to the presence of the cluster tSZ signal and adjacent haloes (both correlated and uncorrelated haloes in the line-of-sight) in the second leg L of the QE. The residuals from the fitting process also contribute to the excess variance but they are sub-dominant compared to the tSZ signals in the L map. The extra variance can be reduced by removing a matched filter estimate of the tSZ from the second leg for the tSZ-free gradient and inpainted gradient QE as demonstrated by Patil *et al.* [60].

3.2.4 Discussion and applicability to real data

In the previous sections, we have shown that removing the cluster SZ signals using the inpainting technique performs better than simple gradient scale cuts in the QE. This is because adopting a gradient scale cut only removes the foreground power from scales smaller than $\ell > \ell_G$ leaving the foregrounds intact on large-scales $\ell \leq \ell_G$. Since the foregrounds have non-zero off-diagonal elements in the harmonic space, unlike CMB, the foreground-induced correlations between the large-scales ($\ell \leq \ell_G$) in the gradient G map and the small-scales in the L map remain unaltered. The level of bias from the residual foregrounds after the gradient cut depends on the foreground power relative to the experimental noise. For example, Baxter *et al.* [30] adopted $\ell_G = 1500$ to reduce tSZ-induced lensing bias to $\sim 10\%$ for SPT-SZ experiment ($\Delta_T = 18\ \mu\mathrm{K}'$) while for SPTpol, which is 4× deeper than SPT-SZ, the tSZ-bias with $\ell_G = 1500$ is $\sim 50\%$ (see Fig. 2 of [R18]). The bias from kSZ was much smaller than the statistical uncertainties in the above SPT analyses while we have shown in this work that it is not true for CMB-S4 datasets. This is consistent with the reduction in level of bias for $\ell_G = 2000$ vs $\ell_G = 2700$ in the sQE: 6% vs 17% for mock kSZ (§3.2.1) and 3% vs 10% for kSZ from S10 simulations (§3.2.2).

By contrary, inpainting completely removes the spatially-localized foreground signals present inside the inpainting radius. The foreground signals from haloes correlated to the cluster outside the inpainting radius can also be important but are highly sub-dominant. The mock kSZ case where the foreground signal is confined to the central $2'$ region clearly illustrates this. It also indicates that the robustness of the inpainted gradient QE does not come from the approximate subtraction of SZ signals outside the $R1$ in inpainted gradient estimator.

The cluster foreground signals used for the tests are extracted from N-body simulations [S10] that mimics the actual data expected from CMB experiments. This ensures that we capture all of the real world effects like, for example, asymmetric cluster SZ profiles, SZ signals from correlated structures outside the cluster radius, offsets between the true cluster center and the SZ centers, scatter in the cluster SZ flux-mass relation, and varied range of cluster velocities and optical depth relevant for the cluster kSZ signals. The inpainting radius $R1 = 4'$ used in this work was chosen conservatively. It encompasses the cluster foreground signals expected from clusters from the next generation CMB surveys [14, 61, 62]. The sizes of these clusters span roughly few arcminutes on the sky. However, we have also shown that $R1$ must be increased slightly when dealing with extremely massive clusters. In general, we find $R1 \sim 2\theta_{200c}$ to be a good choice for all clusters.

4 Conclusion

In this work we have presented a modification to the standard lensing QE by inpainting the large-scale CMB gradient at the cluster location. Using this estimator, we predict that the stacked mass of a cluster sample containing 5000 clusters ($M_{200c} = 2 \times 10^{14}$ $M_\odot$ at $z = 0.7$) can be constrained to $\Delta M/M = 6.5\%$ with a CMB-S4 like low-noise ($\Delta_T = 1.5$ μK′) dataset without performing any foreground cleaning. We compared the estimator to the standard and tSZ-free gradient QE to show that the noise from inpainting is negligible and does not affect the lensing S/N. We studied the effect of cluster SZ biases on all the three estimators using realistic SZ signals from S10 simulations. Our results demonstrate that the inpainted gradient QE is robust against both cluster tSZ and kSZ signals and a similar result cannot be obtained using simple gradient scale cuts in the sQE or the tSZ-free gradient QE. In addition to the SZ-induced lensing biases, we have studied the effect of low mass bias arising from the underestimation of the CMB gradient due to the cluster convergence signal. We find that it is safe to use gradient modes up to $\ell \leq \ell_G$ (= 2700) without worrying about the cluster convergence bias when reconstructing the convergence signal, of clusters with $M_{200c} \lessapprox 4 \times 10^{14}$ $M_\odot$, using the future CMB surveys. By doing so, we achieve a 14% improvement in the mass constraints compared to the fiducial analysis with $\ell_G = 2000$.

The inpainted gradient estimator presented here will be viable to remove SZ biases for the small-scale cluster lensing reconstruction with the CMB temperature data from the near-term and future CMB experiments. The study about the effect of cluster convergence bias performed in this work is also relevant for lensing reconstruction using the polarization data with the future datasets. In addition to the sQE, the inpainting method can also be applied to the tSZ-free gradient and gradient inversion technique [23] for an effective suppression of the kSZ-induced lensing biases.

While not explicitly shown here, the inpainted gradient QE can also be employed to handle the biases due to dust and synchrotron emissions from member galaxies present inside the cluster. But, we leave a careful consideration of these effects to a future work. In addition to the biases considered here, there are other potential systematics that can produce dipole signals just like the cluster lensing. These are due to (a) the transverse motion of the cluster and correlated haloes referred commonly as the moving-lens effect [19, 63, 64]; and (b) the rotational kSZ because of the rotation of the galaxy cluster [19, 25, 65, 66]. Both these effects may extend outside the inpainting radius. However, in both these cases the direction of the dipole is oriented toward the direction of cluster motion or the rotation, while the CMB-cluster lensing dipole is oriented in the direction of the background CMB gradient. Since the background CMB gradient is not correlated to cluster motion or rotation, the moving-lens effect and the rotational kSZ should average down to zero in the stacked measurements presented here.

The inpainting technique is not limited to cluster lensing and, as demonstrated in an earlier work by Benoit-Lévy *et al.* [51], can also be used to fill the masked regions when reconstructing the CMB lensing power spectrum by the LSS. However, to deal with the SZ-induced biases, from unresolved haloes in particular, for the CMB lensing by LSS the methods prescribed by Madhavacheril and Hill [34], Ferraro and Hill [67], Schaan and Ferraro [68] or foreground bias-hardening techniques [69, 70] will yield better results at the cost of a slightly higher lensing reconstruction noise. The tSZ bias can be removed using a tSZ-free gradient gradient QE [34]. The bias from kSZ can be lowered by removing small-scale modes from lensing reconstruction in both legs [67] or just the gradient leg with a degradation in the

lensing S/N. Performing the lensing reconstruction using shear estimators is also another novel technique to remain insensitive to the foregrounds biases [68].

Acknowledgements

We are grateful to Nicholas Battaglia, Thomas Crawford, and Emmanuel Schaan for their constructive feedback on the manuscript. We also thank the anonymous referee for their useful suggestions that helped in shaping this manuscript better.

The UCLA authors acknowledge support from NSF grants AST-1716965 and CSSI-1835865. SP acknowledges support from Melbourne International Engagement Award (MIPP) and Laby Travel Bursary. CR acknowledges the support from Australian Research Council's Discovery Projects scheme (DP150103208). This work was performed in the context of the South Pole Telescope scientific program. SPT is supported by the National Science Foundation through grant PLR-1248097. Partial support is also provided by the NSF Physics Frontier Center grant PHY-1125897 to the Kavli Institute of Cosmological Physics at the University of Chicago, the Kavli Foundation and the Gordon and Betty Moore Foundation grant GBMF 947. A part of this research was carried out at the Jet Propulsion Laboratory, California Institute of Technology, under contract with the National Aeronautics and Space Administration. This research used resources of the National Energy Research Scientific Computing Center (NERSC), a DOE Office of Science User Facility supported by the Office of Science of the U.S. Department of Energy under Contract No. DE-AC02-05CH11231. We acknowledge the use of `HEALPix` [71] and `CAMB` [49] routines.

References

[1] L. E. Bleem, B. Stalder, T. de Haan, K. A. Aird, S. W. Allen, D. E. Applegate, M. L. N. Ashby, M. Bautz, M. Bayliss, *et al.*, The Astrophysical Journal Supplement Series **216**, 27 (2015), arXiv:1409.0850 [astro-ph.CO] .

[2] M. Hilton, M. Hasselfield, C. Sifón, N. Battaglia, S. Aiola, V. Bharadwaj, J. R. Bond, S. K. Choi, D. Crichton, R. Datta, *et al.*, The Astrophysical Journal Supplement Series **235**, 20 (2018), arXiv:1709.05600 [astro-ph.CO] .

[3] E. Rozo, R. H. Wechsler, E. S. Rykoff, J. T. Annis, M. R. Becker, A. E. Evrard, J. A. Frieman, S. M. Hansen, J. Hao, D. E. Johnston, B. P. Koester, T. A. McKay, E. S. Sheldon, and D. H. Weinberg, The Astrophysical Journal **708**, 645 (2010), arXiv:0902.3702 [astro-ph.CO] .

[4] A. Saro, J. J. Mohr, G. Bazin, and K. Dolag, The Astrophysical Journal **772**, 47 (2013), arXiv:1203.5708 [astro-ph.CO] .

[5] J. Ruel, G. Bazin, M. Bayliss, M. Brodwin, R. J. Foley, B. Stalder, K. A. Aird, R. Armstrong, M. L. N. Ashby, M. Bautz, *et al.*, The Astrophysical Journal **792**, 45 (2014), arXiv:1311.4953 [astro-ph.CO] .

[6] S. W. Allen, D. A. Rapetti, R. W. Schmidt, H. Ebeling, R. G. Morris, and A. C. Fabian, MNRAS **383**, 879 (2008), arXiv:0706.0033 .

[7] A. Mantz, S. W. Allen, H. Ebeling, D. Rapetti, and A. Drlica-Wagner, MNRAS **406**, 1773 (2010), arXiv:0909.3099 [astro-ph.CO] .

[8] S. W. Allen, A. E. Evrard, and A. B. Mantz, ARA&A **49**, 409 (2011), arXiv:1103.4829 [astro-ph.CO] .

[9] T. McClintock, T. N. Varga, D. Gruen, E. Rozo, E. S. Rykoff, T. Shin, P. Melchior, J. DeRose, S. Seitz, J. P. Dietrich, *et al.*, MNRAS **482**, 1352 (2019), arXiv:1805.00039 .

[10] H. Miyatake, N. Battaglia, M. Hilton, E. Medezinski, A. J. Nishizawa, S. More, S. Aiola, N. Bahcall, J. R. Bond, E. Calabrese, *et al.*, arXiv e-prints , arXiv:1804.05873 (2018), arXiv:1804.05873 [astro-ph.CO] .

[11] J. P. Dietrich, S. Bocquet, T. Schrabback, D. Applegate, H. Hoekstra, S. Grandis, J. J. Mohr, S. W. Allen, M. B. Bayliss, B. A. Benson, *et al.*, MNRAS **483**, 2871 (2019), arXiv:1711.05344 [astro-ph.CO] .

[12] S. Raghunathan, S. Patil, E. J. Baxter, F. Bianchini, L. E. Bleem, T. M. Crawford, G. P. Holder, A. Manzotti, and C. L. Reichardt, JCAP **8**, 030 (2017), arXiv:1705.00411 .

[13] CMB-S4 Collaboration, K. N. Abazajian, P. Adshead, Z. Ahmed, S. W. Allen, D. Alonso, K. S. Arnold, C. Baccigalupi, J. G. Bartlett, N. Battaglia, *et al.*, ArXiv e-prints (2016), arXiv:1610.02743 .

[14] The Simons Observatory Collaboration, P. Ade, J. Aguirre, Z. Ahmed, S. Aiola, A. Ali, D. Alonso, M. A. Alvarez, K. Arnold, P. Ashton, *et al.*, ArXiv e-prints (2018), arXiv:1808.07445 .

[15] U. Seljak and M. Zaldarriaga, The Astrophysical Journal **538**, 57 (2000), astro-ph/9907254 .

[16] S. Dodelson, Physical Review D. **70**, 023009 (2004), arXiv:astro-ph/0402314 .

[17] G. Holder and A. Kosowsky, The Astrophysical Journal **616**, 8 (2004), astro-ph/0401519 .

[18] M. Maturi, M. Bartelmann, M. Meneghetti, and L. Moscardini, A&A **436**, 37 (2005), astro-ph/0408064 .

[19] A. Lewis and A. Challinor, Phys. Rept. **429**, 1 (2006), arXiv:astro-ph/0601594 .

[20] W. Hu, S. DeDeo, and C. Vale, New Journal of Physics **9**, 441 (2007), arXiv:astro-ph/0701276 .

[21] J. Yoo and M. Zaldarriaga, Physical Review D. **78**, 083002 (2008), arXiv:0805.2155 .

[22] J.-B. Melin and J. G. Bartlett, A&A **578**, A21 (2015), arXiv:1408.5633 .

[23] B. Horowitz, S. Ferraro, and B. D. Sherwin, MNRAS **485**, 3919 (2019).

[24] M. Madhavacheril, N. Sehgal, R. Allison, N. Battaglia, J. R. Bond, E. Calabrese, J. Caliguiri, K. Coughlin, D. Crichton, R. Datta, *et al.*, Physical Review Letters **114**, 151302 (2015), arXiv:1411.7999 .

[25] E. J. Baxter, R. Keisler, S. Dodelson, K. A. Aird, S. W. Allen, M. L. N. Ashby, M. Bautz, M. Bayliss, B. A. Benson, L. E. Bleem, *et al.*, The Astrophysical Journal **806**, 247 (2015), arXiv:1412.7521 .

[26] Planck Collaboration, P. A. R. Ade, N. Aghanim, M. Arnaud, M. Ashdown, J. Aumont, C. Baccigalupi, A. J. Banday, R. B. Barreiro, J. G. Bartlett, *et al.*, A&A **594**, A24 (2016), arXiv:1502.01597 .

[27] G. Hurier and R. E. Angulo, A&A **610**, L4 (2018), arXiv:1711.06029 [astro-ph.CO] .

[28] Í. Zubeldia and A. Challinor, arXiv e-prints , arXiv:1904.07887 (2019), arXiv:1904.07887 [astro-ph.CO] .

[29] J. E. Geach and J. A. Peacock, Nature Astronomy **1**, 795 (2017), arXiv:1707.09369 .

[30] E. J. Baxter, S. Raghunathan, T. M. Crawford, P. Fosalba, Z. Hou, G. P. Holder, Y. Omori, S. Patil, E. Rozo, T. M. C. Abbott, *et al.*, MNRAS (2018), 10.1093/mnras/sty305, arXiv:1708.01360 .

[31] S. Raghunathan, S. Patil, E. Baxter, B. A. Benson, L. E. Bleem, T. L. Chou, T. M. Crawford, G. P. Holder, T. McClintock, C. L. Reichardt, *et al.*, The Astrophysical Journal **872**, 170 (2019), arXiv:1810.10998 [astro-ph.CO] .

[32] J. E. Geach, J. A. Peacock, A. D. Myers, R. C. Hickox, M. C. Burchard, and M. L. Jones, arXiv e-prints , arXiv:1902.06955 (2019), arXiv:1902.06955 [astro-ph.GA] .

[33] R. A. Sunyaev and Y. B. Zel'dovich, Comments on Astrophysics and Space Physics **2**, 66 (1970).

[34] M. S. Madhavacheril and J. C. Hill, Physical Review D. , 023534 (2018), arXiv:1802.08230 .

[35] R. A. Sunyaev and Y. B. Zeldovich, MNRAS **190**, 413 (1980).

[36] N. Hand, G. E. Addison, E. Aubourg, N. Battaglia, E. S. Battistelli, D. Bizyaev, J. R. Bond, H. Brewington, J. Brinkmann, B. R. Brown, *et al.*, Physical Review Letters **109**, 041101 (2012).

[37] T. Mroczkowski, S. Dicker, J. Sayers, E. D. Reese, B. Mason, N. Czakon, C. Romero, A. Young, M. Devlin, S. Golwala, *et al.*, The Astrophysical Journal **761**, 47 (2012), arXiv:1205.0052 [astro-ph.CO] .

[38] J. Sayers, T. Mroczkowski, M. Zemcov, P. M. Korngut, J. Bock, E. Bulbul, N. G. Czakon, E. Egami, S. R. Golwala, P. M. Koch, *et al.*, The Astrophysical Journal **778**, 52 (2013), arXiv:1312.3680 [astro-ph.CO] .

[39] J. C. Hill, S. Ferraro, N. Battaglia, J. Liu, and D. N. Spergel, Physical Review Letters **117**, 051301 (2016), arXiv:1603.01608 [astro-ph.CO] .

[40] E. Schaan, S. Ferraro, M. Vargas-Magaña, K. M. Smith, S. Ho, S. Aiola, N. Battaglia, J. R. Bond, F. De Bernardis, E. Calabrese, *et al.*, Physical Review D. **93**, 082002 (2016), arXiv:1510.06442 .

[41] B. Soergel, S. Flender, K. T. Story, L. Bleem, T. Giannantonio, G. Efstathiou, E. Rykoff, B. A. Benson, T. Crawford, S. Dodelson, *et al.*, MNRAS **461**, 3172 (2016), arXiv:1603.03904 .

[42] R. Adam, I. Bartalucci, G. W. Pratt, P. Ade, P. André, M. Arnaud, A. Beelen, A. Benoît, A. Bideaud, N. Billot, *et al.*, A&A **598**, A115 (2017), arXiv:1606.07721 [astro-ph.CO] .

[43] F. De Bernardis, S. Aiola, E. M. Vavagiakis, N. Battaglia, M. D. Niemack, J. Beall, D. T. Becker, J. R. Bond, E. Calabrese, H. Cho, *et al.*, JCAP **3**, 008 (2017), arXiv:1607.02139 .

[44] A. Hall and A. Challinor, Physical Review D. **90**, 063518 (2014), arXiv:1407.5135 [astro-ph.CO] .

[45] S. Yasini and E. Pierpaoli, Physical Review D. **94**, 023513 (2016), arXiv:1605.02111 [astro-ph.CO] .

[46] N. Sehgal, P. Bode, S. Das, C. Hernandez-Monteagudo, K. Huffenberger, Y. Lin, J. P. Ostriker, and H. Trac, The Astrophysical Journal **709**, 920 (2010), arXiv:0908.0540 .

[47] Planck Collaboration, P. A. R. Ade, N. Aghanim, M. Arnaud, M. Ashdown, J. Aumont, C. Baccigalupi, A. J. Banday, R. B. Barreiro, J. G. Bartlett, and et al., A&A **594**, A13 (2016), arXiv:1502.01589 .

[48] M. Zaldarriaga and U. Seljak, Physical Review D. **59**, 123507 (1999).

[49] A. Lewis, A. Challinor, and A. Lasenby, The Astrophysical Journal **538**, 473 (2000).

[50] J. Silk, The Astrophysical Journal **151**, 459 (1968).

[51] A. Benoit-Lévy, T. Déchelette, K. Benabed, J. F. Cardoso, D. Hanson, and S. Prunet, A&A **555**, A37 (2013), arXiv:1301.4145 [astro-ph.CO] .

[52] E. M. George, C. L. Reichardt, K. A. Aird, B. A. Benson, L. E. Bleem, J. E. Carlstrom, C. L. Chang, H.-M. Cho, T. M. Crawford, A. T. Crites, *et al.*, The Astrophysical Journal **799**, 177 (2015), arXiv:1408.3161 .

[53] J. F. Navarro, C. S. Frenk, and S. D. M. White, The Astrophysical Journal **462**, 563 (1996), arXiv:astro-ph/9508025 .

[54] M. Bartelmann, A&A **313**, 697 (1996), arXiv:astro-ph/9602053 .

[55] A. R. Duffy, J. Schaye, S. T. Kay, and C. Dalla Vecchia, MNRAS **390**, L64 (2008), arXiv:0804.2486 .

[56] Planck Collaboration, P. A. R. Ade, N. Aghanim, M. Arnaud, M. Ashdown, J. Aumont, C. Baccigalupi, A. Balbi, A. J. Banday, R. B. Barreiro, E. Battaner, *et al.*, A&A **561**, A97 (2014), arXiv:1303.5090 [astro-ph.CO] .

[57] D. Alonso, T. Louis, P. Bull, and P. G. Ferreira, Physical Review D. **94**, 043522 (2016), arXiv:1604.01382 [astro-ph.CO] .

[58] V. Vikram, A. Lidz, and B. Jain, MNRAS **467**, 2315 (2017), arXiv:1608.04160 [astro-ph.CO] .

[59] J. C. Hill, E. J. Baxter, A. Lidz, J. P. Greco, and B. Jain, Physical Review D. **97**, 083501 (2018), arXiv:1706.03753 [astro-ph.CO] .

[60] S. Patil, S. Raghunathan, and C. L. Reichardt, arXiv e-prints , arXiv:1905.07943 (2019), arXiv:1905.07943 [astro-ph.CO] .

[61] B. A. Benson, P. A. R. Ade, Z. Ahmed, S. W. Allen, K. Arnold, J. E. Austermann, A. N. Bender, L. E. Bleem, J. E. Carlstrom, C. L. Chang, *et al.*, in *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, Vol. 9153 (2014) p. 1, arXiv:1407.2973 [astro-ph.IM] .

[62] K. Abazajian, G. Addison, P. Adshead, Z. Ahmed, S. W. Allen, D. Alonso, M. Alvarez, A. Anderson, K. S. Arnold, C. Baccigalupi, *et al.*, arXiv e-prints , arXiv:1907.04473 (2019), arXiv:1907.04473 [astro-ph.IM] .

[63] S. C. Hotinli, J. Meyers, N. Dalal, A. H. Jaffe, M. C. Johnson, J. B. Mertens, M. Münchmeyer, K. M. Smith, and A. van Engelen, arXiv e-prints , arXiv:1812.03167 (2018), arXiv:1812.03167 [astro-ph.CO] .

[64] S. Yasini, N. Mirzatuny, and E. Pierpaoli, The Astrophysical Journal Letters **873**, L23 (2019), arXiv:1812.04241 [astro-ph.CO] .

[65] A. S. Baldi, M. De Petris, F. Sembolini, G. Yepes, W. Cui, and L. Lamagna, MNRAS **479**, 4028 (2018), arXiv:1805.07142 [astro-ph.CO] .

[66] E. J. Baxter, B. D. Sherwin, and S. Raghunathan, arXiv e-prints (2019), arXiv:1904.04199 [astro-ph.CO] .

[67] S. Ferraro and J. C. Hill, Physical Review D. **97**, 023512 (2018), arXiv:1705.06751 [astro-ph.CO] .

[68] E. Schaan and S. Ferraro, arXiv e-prints , arXiv:1804.06403 (2018), arXiv:1804.06403 [astro-ph.CO] .

[69] T. Namikawa, D. Hanson, and R. Takahashi, MNRAS **431**, 609 (2013), arXiv:1209.0091 [astro-ph.CO] .

[70] S. J. Osborne, D. Hanson, and O. Doré, JCAP **2014**, 024 (2014), arXiv:1310.7547 [astro-ph.CO] .

[71] K. M. Górski, E. Hivon, A. J. Banday, B. D. Wandelt, F. K. Hansen, M. Reinecke, and M. Bartelmann, The Astrophysical Journal **622**, 759 (2005), arXiv:astro-ph/0409513 .